We claim:

1. A method executed by a computer under the control of a program, said computer including a memory for storing said program, said method comprising the steps of:

a) inputting an ensemble of protein backbone scaffolds;

b) applying at least one protein design cycle to each of said scaffolds; and

c) generating a probability matrix derived from a plurality of variable sequences.

2. A method according to claim 1 wherein said protein design cycle comprises a sequence prediction algorithm.

3. A method according to claim 1 wherein said protein design cycle comprises a dead end elimination algorithm.

4. A method according to claim 1 wherein said protein design cycle comprises a genetic algorithm.

5. A method according to claim 1 wherein said protein design cycle comprises a Monte Carlo algorithm.

6. A method according to claim 1 wherein said protein design cycle comprises a self consistent mean field theory (SCMF) algorithm.

7. A method according to claim 1 wherein said protein design cycle comprises a Monte Carlo algorithm.

8. A method according to claim 1 further comprising ranking said variable sequences.

9. A method according to claim 1 further comprising synthesizing a plurality of said variable sequences.

10. A method according to claim 1 further comprising recombining a plurality of said variable sequences to form additional variable sequences.

11. A method according to claim 1 wherein said ensemble comprises a family of naturally occurring proteins.

12. A method according to claim 1 wherein said ensemble is generated by a Monte Carlo simulation.

13. A method according to claim 1 wherein said ensemble is generated by a Molecular Dynamics simulation.

14. A method according to claim 1 wherein said ensemble is derived from an NMR ensemble of structures.

15. A method according to claim 1 wherein after step a), a rotamer library is pre-filtered to eliminate high energy interactions to form a suitable rotamer set for each scaffold.

16. A method for optimizing simulation or scoring function parameters that utilizes comparisons between designed sequences and natural sequences, comprising the steps of:

a) applying a protein design cycle to produce a variable protein sequence;

b) comparing said variable protein sequence to at least one natural protein

sequence and/or conformation;

c) modifying said simulation or scoring function parameters to reflect said comparison.

17. A method for optimizing simulation or scoring function parameters that utilizes comparisons between designed sequences and natural sequences, comprising the steps of:

a) applying a protein design cycle to produce an amino acid probability matrix;

b) comparing said matrix to at least one natural protein sequence and/or conformation;

c) modifying said simulation or scoring function parameters to reflect said comparison.